

ПРАВИТЕЛЬСТВО РОССИЙСКОЙ ФЕДЕРАЦИИ

**Федеральное государственное автономное
образовательное учреждение высшего образования**

**Национальный исследовательский университет
«Высшая школа экономики»**

Факультет экономических наук
Образовательная программа «Экономика»

КУРСОВАЯ РАБОТА

На тему «Вариация алгоритма кросс-валидации со взвешиванием наблюдений»

Студент группы БЭК161
Гармидер Петр Александрович

Научный руководитель:
Борис Демешев

Москва 2019

Содержание

1	Введение	3
1.1	Параметры и гиперпараметры модели	3
1.2	Кросс-Валидация	3

1 Введение

1.1 Параметры и гиперпараметры модели

Большая часть популярных моделей имеют множество параметров и гиперпараметров однозначно выделяющие модель из множества всех алгоритмов. Гиперпараметры модели — параметры, вводимые пользователем вручную, которые в большинстве случаев не меняются в ходе обучения¹. Параметры модели — параметры, которые не вводятся пользователем вручную, а есть результат оптимизации функции потерь. Их конечные значения становятся известны после завершения процесса обучения. Параметры модели связывают имеющиеся у исследователя данные и выбранный алгоритм для обучения.

Рассмотрим пример. Пусть имеем обучающую выборку $D = \{x_i, y_i\}$ и тестовую выборку $d = \{x_i, y_i\}$, где $x_i, y_i \in R$, $|D| = N$ и $|d| = n$. Будем оценивать y_i используя Ridge-regression (см. [2]). Имеем безусловную задачу оптимизации:

$$\sum_{i=1}^N (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2 + \lambda \hat{\beta}_1^2 \rightarrow \min_{\hat{\beta}_0, \hat{\beta}_1}$$

В (1.1) существуют два параметра для оптимизации $\hat{\beta}_0$ и $\hat{\beta}_1$, а также один гиперпараметр λ , который является константой в рассматриваемой задаче. Коэффициенты регрессии находятся из задачи минимизации при заданном уровне λ , в то время, как степень регуляризации λ задается исследователем. В конечном счете исследователь заинтересован в создании модели, которая имеет высокое качество на тестовой выборке d , что не участвовала в процессе обучения. Качество работы модели на выборке d очевидно зависит от выбранного исследователем λ . Поэтому, выбор оптимального значения для гиперпараметров является также важной задачей для получения лучшей модели.

Гиперпараметрами модели также могут быть: метод обработки пропусков в данных, количество слоев в нейронной сети, выбранные функции активации, уровень Dropout, скорость обучения и другие.

Влияние гиперпараметров на качество работы модели делает их правильный выбор отдельной задачей. Оптимальный алгоритм подбора гиперпараметров уникален для каждой конкретной задачи. Качество работы модели в целом принято оценивать на выборках не участвовавших в обучении, но для которых известно истинное значение зависимой переменной.

1.2 Кросс-валидация

¹Однако такая практика также используется, например, изменение гиперпараметра learning rate в ходе алгоритма градиентного спуска (см. [1])

Список литературы

- [1] Zeiler Matthew D. ADADELTA: An Adaptive Learning Rate Method. 2012.
- [2] Hoerl Arthur E, Kennard Robert W. Ridge regression: Biased estimation for nonorthogonal problems // Technometrics. 1970. T. 12, № 1. С. 55–67.